心理科学进展 2018, Vol. 26, No. 6, 1031–1040
Advances in Psychological Science
DOI: 10.3724/SP.J.1042.2018.01031

孤独症谱系障碍者的视听时间整合*

李涛涛[1] 胡金生[1] 王 琦[1] 李骋诗[1] 李松泽[1]
何建青[1] 李辰洋[1] 刘淑清[2]
([1]辽宁师范大学心理学院, 大连 116029) ([2]大连医科大学基础医学院, 大连 116044)

摘 要 视听时间整合是指个体对一定时间间隔内输入的视听刺激进行表征的过程，是视听整合的重要机制。孤独症谱系障碍者的视听时间整合存在缺陷，主要表现为4个方面：视听时间整合窗口较正常个体更宽、更对称；快速视听时间再校准能力不足；听觉时间线索对其视觉搜索促进作用弱；言语刺激的视听时序知觉敏感性低。目前使用的研究任务多样，如声音诱发闪光错觉和"pip-pop"任务从内隐角度探究视听整合中的时间机制，同时性判断、时序判断和优先注视任务主要用于跨通道时序知觉的研究。相关理论从神经加工异常、先验经验不足和视听通道的相互作用等角度解释了其缺陷。未来需要进一步提高研究生态效度，整合理论解释，精确量化诊断指标，同时开发有实效的干预策略。

关键词 孤独症谱系障碍；视听时间整合；时间窗；时间同步性
分类号 B842; R395

1 引言

个体对外界事物的感知依赖于多感觉通道信息的整合。例如，在言语交流时，需要整合谈话者的表情、手势和声音等信息。视听时间整合(audiovisual temporal integration)是个体对一定时间间隔内输入的视听刺激进行表征的过程，即对时间上不同步的刺激进行再加工，将其整合为一个有意义的独立事件(Stevenson, Fister, Barnett, Nidiffer, & Wallace, 2012; Stevenson & Wallace, 2013)。整合的时间窗口模型(Time-Window-of-Integration)认为，视觉、听觉和触觉等感觉线索是并行输入的，其加工过程及所用加工时间彼此独立，一旦某个通道的刺激完成加工，整合窗口便会开启，如果其他通道的信息输入也在该窗口内终止，多感觉整合就会发生(Colonius & Diederich, 2004)。

孤独症谱系障碍(Autism Spectrum Disorders, ASD)是一组广泛性神经发育障碍，感知觉异常是其症状的主要方面，并已作为核心诊断标准纳入了美国精神障碍诊断统计手册第五版(DSM-V) (American Psychiatric Association, 2013)。早期对ASD者感知觉缺陷的研究多关注单通道信息加工的异常。例如，ASD者对音高、音强等低级声学线索过度敏感，会出现听觉解码偏差(O'Connor, 2012)。近年来，跨通道感觉的加工模式则倍受关注，神经生理学的研究发现，ASD者在进行视听整合时与正常个体的神经激活模式存在差异，如有关功能性磁共振成像(functional magnetic resonance imaging, fMRI)的研究表明，视听情绪信息整合时，控制组会激活额颞叶相关皮质区，而ASD者更多激活顶额网络(Doyle-Thomas, Goldberg, Szatmari, & Hall, 2013)；关于视听语义整合时间进程的研究表明，在时间窗为200~300 ms时，听觉会促进视觉唇动加工，而ASD者在该时段内两者间的调节作用不明显(Megnin et al., 2012)。同样，行为研究结果表明，ASD者在整合视听、视触和听触等信息时均存在缺陷，主要表现为对复杂的社会性刺激难以整合(Baum, Stevenson, & Wallace, 2015; Greenfield, Ropar, Smith, Carey, & Newport, 2015; Russo et al., 2010)。根据弱中央统合(weak central coherence)理论，"统合"

收稿日期：2017-04-06
* 辽宁省特聘教授基金、辽宁省人文社科重点研究基地项目和辽宁省儿童青少年健康人格评定与培养协同创新中心资助项目支持
通信作者：胡金生, E-mail: hu_jinsheng@126.com

需要将信息整合到相关的情境中来完成更高水平的心理完形，该过程不仅需要注意资源的投入，也依赖于对信息时空特性的把握，跨通道刺激间时间线索加工的低敏感性是整合加工缺陷产生的重要原因(Glessner et al., 2009; Stevenson, Zemtsov, & Wallace, 2012)。

ASD 者视听时间整合的研究可以进一步探明其跨通道整合加工缺陷的形成机制，相关研究有利于超越其感知觉缺陷表型的多维性，为揭示其病因的复杂性和相通性提供理论基础。例如，时间窗口作为视听时间整合的重要衡量指标，能够直观反映 ASD 者和正常被试在短时间内对视听刺激的整合加工模式差异。研究表明，ASD 者时间整合窗口比正常人宽，同等条件下，该群体对跨通道感觉刺激输入的时间估计存在偏差，分离异步性刺激的能力较弱(Wallace & Stevenson, 2014)。另外，明确该群体视听整合的时间关系，不仅有利于在后续的行为和知觉训练中对感知刺激的时间知觉进行精准干预，还有利于说明 ASD 者言语障碍、社交缺陷等核心症状的形成机制。例如，唇动和声音线索之间的时间同步性是视听言语整合的关键，ASD 者对视听同步性辨别敏感性的降低会导致其言语接受能力减弱(Patten, Watson, & Baranek, 2014)。

2 孤独症者视听时间整合的研究方法

2.1 声音诱发闪光错觉任务

声音诱发闪光错觉任务(sound-induced flash illusion task)主要从内隐的角度探讨 ASD 者的视听时间整合。在该任务中，呈现一个闪光和两个声音，其中一个声音与闪光同时呈现，另一个声音与闪光不同时呈现。要求被试忽略听觉刺激，报告看到的闪光数量(Foss-Feig et al., 2010)。在此过程中，闪光和声音之间的时间关系会对闪光数量的知觉产生影响，当时间间隔变小或者对其时间关系的知觉能力降低时，闪光数量知觉的错误率会增加(Shams, Kamitani, & Shimojo, 2002)。该任务主要使用非言语刺激进行研究，实验中同时呈现视听刺激，这与现实生活很相符，因此具有较高的生态效度。

2.2 “pip-pop”任务

“pip-pop”任务是通过呈现一系列方向不同、且在两种颜色间随机变化的短线条，其中水平或垂直的线条为靶刺激，在靶刺激颜色变化的同时呈现一个声音。要求被试搜索靶刺激，并记录其搜索的正确率和反应时。该任务用于探讨视听整合中时间加工机制的原理是：视觉搜索中，在一定的时间内呈现声音时可以促进搜索效率的提高，即产生“pop-out”效应。当视听刺激间的时间知觉能力减弱时，该效应会随之减弱(van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008; Collignon et al., 2013)。该任务从一个新的角度对 ASD 者的视听时间整合进行了研究，但需要被试进行按键反应，因此只适用于高功能 ASD 个体。另外，实验要求记录反应时，而 ASD 者普遍存在注意缺陷，因此其可靠性还需进一步提高。

2.3 同时性判断任务

同时性判断任务(simultaneity judgment task, SJ)主要用于检测 ASD 者视听时间同步性辨别能力。在该任务中，先呈现视觉或听觉通道刺激，一个时间间隔(stimulus onset asynchrony, SOA)后，再呈现另一通道的刺激。要求被试判断两者是否同时出现。通过操控 SOA 来测量同步性知觉敏感性，研究中根据需要可设置不同的 SOA，一般非言语刺激为 0~300 ms，言语刺激为 0~400 ms (Stevenson, Siemann, Schneider et al., 2014)。根据正确率和 SOA 值拟合函数，得到时间窗和主观同时点(Baum et al., 2015)。该任务设计巧妙，易于操作，但需要被试听懂指导语并按键反应，因此只适用于高功能被试。不过，即使是高功能 ASD 者也存在言语理解障碍，在研究中是否听懂指导语还需谨慎确认。

2.4 时序判断任务

时序判断任务(temporal order judgment task, TOJ)是用于测量时序辨别敏感性的方法，依据反应通道的不同，可分为视觉、听觉和多感觉 TOJ 任务，视听时间整合的研究通常使用多感觉 TOJ 任务。该任务的操作方法和 SJ 任务基本相同，但要求被试判断的是哪个刺激先出现。实验中，一般视觉先于听觉刺激(visual preceding auditory, VA)和听觉先于视觉刺激(auditory preceding visual, AV)呈现的概率各占 50%。时序知觉的敏感性通常根据主观同时点和最小可觉差(just noticeable differences, JND)来衡量(de Boer-Schellekens, Eussen, & Vroomen, 2013; Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011)。相比上

文提到的同时性判断任务，该任务较少受反应偏差的影响，精准性更好，但需要投入更多地认知加工，对被试也更具挑战性(Binder, 2015; Vroomen & Keetels, 2010)。

2.5 优先注视任务

优先注视任务(preferential-looking task)是基于眼动技术提出的，主要用于检验 ASD 者对言语刺激的时间同步性知觉。在该任务中，呈现两个不同步但内容相同的视频，同时通过耳机播放与视频相匹配的声音，它只与其中一个视频保持同步。其原理为：当视听信息之间存在固定的时间关系时，听觉信息会影响视觉的注视模式。当声音的呈现和两个视频的同步性随机匹配时，被试对每个视频的注视时间应分别为 50%；而当声音呈现与视频之间的时间关系固定时，个体对视听同步呈现视频的注视时间会更长(Bebko, Weiss, Demark, & Gomez, 2006; Grossman, Steinhart, Mitchell, & McIlvane, 2015)。该任务指导语简单，对被试的言语和认知能力要求较低，而且不需要被试按键反应，数据搜集方式无干扰，实验过程趣味性较强，适用于幼儿和低功能被试。

从上文的研究方法可以看出，ASD 者视听时间整合的研究任务多样，测量角度颇为丰富。声音诱发闪光错觉和“pip-pop”任务主要从内隐角度探讨视听整合中，听觉对视觉加工直接调节作用的时间点和时间范围；优先注视任务测量 ASD 和正常被试在视听同步性知觉中的眼动加工差异；SJ 和 TOJ 任务从外显角度测量跨通道时序知觉的敏感性，任务中的主要变量为刺激呈现的同步和异步，前者只需知觉时间顺序，后者还需辨别顺序的先后，因此需要更多的心理加工。研究发现，ASD 者对不同刺激属性的视听时间知觉受任务类型的影响，如对于非言语刺激，该群体在声音诱发闪光错觉任务中表现出时间窗宽的特点，但在优先注视任务、SJ 和 TOJ 任务中却与正常被试无显著差异(Foss-Feig et al., 2010; Bebko et al., 2006; Stevenson, Siemann, Schneider et al., 2014)。这表明 ASD 者跨通道时序知觉和内隐视听时间加工可能存在不同的加工机制。

3 孤独症者视听时间整合缺陷表现

3.1 视听时间整合窗口异常

视听整合过程中，个体的知觉能够忽略视听刺激之间可能存在的时间差，把不同步的刺激整合为一个有意义的整体事件，视听刺激之间的时间间隔被称为时间窗(Temporal Binding Window, TBW)。在该窗口内输入的刺激可以被整合为一个整体，当超出这个时间间隔时，则会被知觉为两个不同事件。它反应了大脑对视听刺激时间间隔知觉敏感性的高低，TBW 越宽表明对刺激间时间差知觉的敏感性越低。当 TBW 变宽时，视听刺激之间的时间关系会变得模糊，该时间范围内输入的刺激可能被整合，但往往对不同事件却不能区别，致使多种感觉的知觉表征减弱，甚至出现错误整合(Stevenson et al., 2016)。很多研究表明，与正常被试相比，ASD 者存在 TBW 宽的特点(Foss-Feig et al., 2010; Kwakye et al., 2011; Stevenson, Siemann, Schneider et al., 2014; Stevenson, Siemann, Woynaroski et al., 2014) (如图 1)。

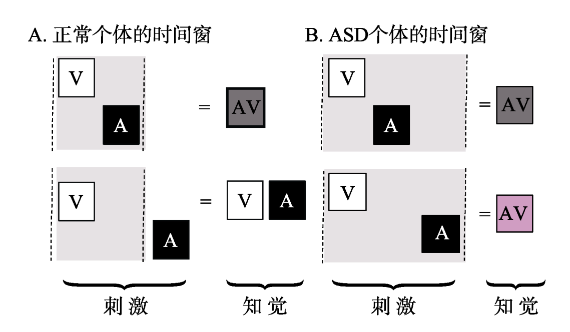


图 1 正常人和 ASD 者的时间窗结构
(资料来源：Stevenson et al., 2016)

首先, ASD 者基于非言语刺激的 TBW 约为正常被试的 2 倍。研究者使用声音诱发闪光错觉任务对 46 名 8~17 岁的 ASD 和正常被试进行了研究。结果发现，当 SOA 在 300 ms (–150 ~ +150 ms)内时，正常组报告两个闪光的概率会提高；而 ASD 被试大约在 600 ms (–300 ~ +300 ms)内都会提高。在基线条件下，即呈现一个圆点和一个声音时，组间差异不显著，这表明两组差异不是由反应偏差造成的，ASD 者的听觉会在更长时段内干扰其视觉加工(Foss-Feig et al., 2010)。随后，研究者直接检验了 ASD 者听觉对视觉时序判断的影响。该研究在视觉 TOJ 任务的基础上，分别在两个视觉刺激之前和之后呈现声音，第一个声音与第一个视觉刺激同时呈现，第二个声音在第二个视觉刺激出现一个 SOA 后呈现。结果发现，正

常被试在 SOA 为 50~150 ms 时声音会促进视觉时序判断, 而 ASD 被试在 0~300 ms 内时序判断的正确率都提高了(Kwakye et al., 2011)。

其次, 刺激类型会影响 TBW 的宽度, ASD 者基于言语刺激的 TBW 较正常被试更宽, 且与麦格克效应的强度有关。该效应主要用于衡量视听言语整合能力, 其强度越强, 视听言语整合能力越弱(Stevenson, Siemann, Schneider et al., 2014; Baum et al., 2015)。一方面, 行为研究检验了 ASD 者时间加工与视听言语整合的关系, 研究中使用简单、复杂非言语和言语三种属性刺激。结果发现, 相比正常被试, ASD 者基于言语刺激的 TBW 更宽, 且该群体麦格克效应的强度从简单非言语刺激、复杂非言语刺激到复杂言语刺激逐渐增强, 而正常组没有表现出这种趋势(Stevenson, Siemann, Schneider et al., 2014; Stevenson, Siemann, Woynaroski et al., 2014b)。另一方面, 神经生理学的研究表明, 后侧颞上沟(Posterior Superior Temporal Sulcus, pSTS)是视听时间和言语整合的重要神经机制, 在麦格克效应中有着重要作用, 而 ASD 者该脑区的激活模式和正常被试存在差异(Hocking & Price, 2008; Stevenson, VanDerKlok, Pisoni, & James, 2011; Beauchamp, Nath, & Pasalar, 2010)。有研究使用 fMRI 技术检验了 ASD 和正常被试对不同属性刺激在视听同步和异步呈现时的神经激活模式, 结果显示, 在言语刺激条件下, 当视听同步呈现时, 正常被试 pSTS 区域的激活显著减少, 而 ASD 者没有显著变化; 在非言语刺激条件下, 视听同步和异步呈现时被试间差异不显著。这表明在视听言语刺激同步呈现时, ASD 者的整合效率没有提高, 其与言语刺激有关的神经加工机制存在异常(Stevenson et al., 2013)。

再次, 言语阅读障碍者也表现出了 TBW 宽的特点, 但与 ASD 者有所不同, 言语阅读障碍者在视觉和听觉时间加工方面都存在缺陷, 而 ASD 者只存在听觉时间加工缺陷(Laasonen, Service, & Virsu, 2001; Hairston, Burdette, Flowers, Wood, & Wallace, 2005; Kwakye et al., 2011)。研究发现, 两者的神经加工模式不同, ASD 者表现为局部超联结和长距联结不足, 而言语阅读障碍者表现为近距离联结不足和末端联结过度(Rippon, Brock, Brown, & Boucher, 2007; Williams & Casanova, 2010)。

最后, ASD 者 TBW 的分布结构更对称。目前, 对于 ASD 者 TBW 的研究主要基于实验室实验, 视听刺激之间的时间间隔受人为控制, 通常 AV 和 VA 两种条件随机呈现。然而, 由于视听信息的传播速度不同, 个体接收到听觉信息的时间应晚于视觉信息, 因此 TBW 为左宽右窄的不对称分布, 而 ASD 个体的 TBW 则更对称, 其分布形态似正常幼儿期(Hillock-Dunn & Wallace, 2012; Hillock, Powers, & Wallace, 2011)。

3.2 快速视听时间再校准能力不足

快速视听时间再校准(rapid audiovisual temporal recalibration)是大脑对视听刺激的时间关系再适应的结果, 当视听不同步时, 可以通过调节两者间时间关系向延迟的方向偏移, 将其整合为一个整体事件(van der Burg, Alais, & Cass, 2013)。校准能力的强弱通常用主观同时点(point of subjective simultaneity, PSS)的偏移进行衡量。PSS 是指主观上认为两个刺激同时出现, 但实际却存在一定的时间间隔(Noel, De Niear, Stevenson, Alais, & Wallace, 2017)。当视听刺激绝对同时呈现时, PSS 应是 SOA 为 0 时的值, 但在进行再校准时, PSS 会发生偏移。其偏移量越大, 表明校准能力越强, 即对视听不同步的适应能力越强。相比正常被试, ASD 者的快速视听时间再校准能力不足表现为不能快速适应视听刺激的不同步性, 且对不同刺激类型再校准能力的共性低(Turi, Karaminis, Pellicano, & Burr, 2016; Noel et al., 2017)。

研究者使用简单非言语刺激对 ASD 和正常成人的快速视听时间再校准进行了研究, 结果发现, 当先呈现听觉先于视觉的试次, 再呈现视觉先于听觉的试次时, 先前试次会对当前试次产生影响。表现为正常成人的 PSS 偏移了 53 ms, 而 ASD 成人几乎没有发生偏移(Turi et al., 2016)。这表明当视听刺激不同步呈现时, 正常成人对优先呈现通道时序知觉的适应阈限发生了变化, 当视听刺激之间的时间间隔变大时, 其适应效应增强, 而 ASD 者不受影响。有研究还考察了 ASD 青少年的快速视听时间再校准与刺激属性的关系。研究中使用了简单非言语、复杂非言语和言语三种刺激。结果表明, 在言语刺激条件下, ASD 被试的再校准能力与正常组无差异, 但在简单和复杂非言语刺激条件下, ASD 被试 PSS 的偏移量显著少于正常组, 简单非言语条件下两组被试的偏移量

分别为 33.15 ms 和 17.96 ms, 复杂非言语条件下分别为 18.06 ms 和 9.12 ms (Noel et al., 2017)。

值得注意的是, 有关正常成人的研究表明, 快速视听时间再校准能力与 TBW 存在正相关, 但在 Noel 等(2017)的研究中却没有发现这种相关(van der Burg et al., 2013)。这可能与年龄因素有关, 前者以成人为研究对象, 而后者则以青少年为被试。贝叶斯理论认为感觉整合的各个方面都会随着年龄的变化而变化, 且正常人的 TBW 到成年期才达到成熟(Lewkowicz & Flom, 2014; Hillock-Dunn & Wallace, 2012)。

3.3　听觉时间线索对视觉搜索促进作用弱

在视听整合过程中, 当视听刺激在一定的时间间隔内呈现时, 大脑会协调视听通道的信息加工使其相互影响。在此过程中, 听觉能够依赖视听刺激间的时间关联性调节视觉通道的加工(Shi, Chen, & Müller, 2010)。"Pop-out"效应是视听通道信息调节作用的一种, 该效应的发生需要对视听信息之间的时间关系进行加工(van der Burg et al., 2008)。有学者对 14~31 岁的 ASD 者和正常被试进行了研究, 研究中通过改变闪光数量设置了三种负荷水平, 分别为 24、36 和 48。结果发现, 当呈现"pip"声音时, 正常被试在所有水平上的反应时和正确率都显著提高, 而 ASD 者均没有提高, 且在没有声音条件下, ASD 被试的反应时快于正常被试(Collignon et al., 2013)。这表明 ASD 者的视听整合缺陷可能是由于其不能精确知觉视听刺激间的时间关系造成的。

然而, 有研究却得出了不同的结论。de Boer-Schellekens 和 Keetels 等(2013)使用三种任务进行了研究: 伴随声音的视觉时序判断任务、"pip-pop"任务、数字表阅读任务。结果发现, ASD 个体视觉时序辨别敏感性较低, 但当呈现声音刺激时, ASD 和正常被试的 JND 都显著降低, 即敏感性提高了, 且 ASD 被试 JND 下降的幅度大于正常被试。这表明 ASD 个体可以利用声音线索促进视觉时序判断。同样, 在"pip-pop"任务中 当呈现声音时, 与正常被试一样, ASD 者的视觉搜索效率提高了, 这说明 ASD 个体能够知觉视听刺激的时间同步性。

结果的争议可能是由于 ASD 者的注意缺陷造成的, 研究发现分散的视觉注意有利于"pip-pop"任务中的视听同步性知觉(van der Burg et al., 2008)。在 de Boer-Schellekens 和 Keetels 等(2013)的研究中直接告诉被试声音和靶刺激颜色的变化同时呈现, 这使得 ASD 被试可以自由使用搜索策略, 利用自上而下的认知加工促进视觉搜索, 弥补了自下而上的视听整合缺陷。而在 Collignon 等(2013)的研究中则要求被试注视中心注视点, 由于 ASD 者弱中央统合的加工特点, 会强化其局部加工缺陷, 不利于促进视听信息的同步性知觉(Happé & Frith, 2006)。

3.4　言语刺激的视听时序知觉敏感性低

视听时序知觉是指对两个视听事件是否同时出现以及相继出现时先后顺序的知觉, 其敏感性受刺激属性的影响, 通常, 对言语刺激时序知觉的敏感性低于非言语刺激(Vatakis & Spence, 2006)。多数研究发现, 与正常被试相比, ASD 者的时序辨别敏感性低, 且在刺激属性和任务差异方面存在不同(Grossman et al., 2015; de Boer-Schellekens, Eussen, et al., 2013)。

一方面, 在视听同步性知觉的研究中, ASD 者对言语刺激的视听同步性辨别能力较低。学者使用优先注视任务检验了 4~6 岁的 ASD、正常和其它发育障碍儿童在非言语、简单言语和复杂言语刺激方面的视听同步性辨别能力。结果发现, ASD 儿童基于非言语刺激的时间同步性辨别与正常被试差异不显著, 而在简单和复杂言语刺激方面却显著低于正常组(Bebko et al., 2006)。随后, 有研究使用该任务进一步从内隐和外显两个角度进行了探讨, 结果发现, 控制组在外显条件下对视听同步呈现视频注视的概率高于内隐条件, 而 ASD 组在两种条件下差异不显著, 且注视概率都显著低于控制组(Grossman et al., 2015)。另一方面, 在视听时序判断任务中, ASD 者时序辨别敏感性低但不受刺激类型的影响。研究者探讨了 ASD 者和正常青少年对不同刺激属性的时序知觉, 结果发现, ASD 者的 JND (116.6 ms)大于正常被试(88.1 ms), 但基于刺激属性的组间差异不显著(de Boer-Schellekens, Eussen, et al., 2013)。

综上所述, ASD 者基于言语刺激的时间同步性辨别存在缺陷。一种可能是由于该群体对言语事件之间的时间同步性尚未建立; 另一种可能是实验中试次呈现的时间太短, 其辨别能力弱致使没有觉察到视听刺激的同步性。此外, ASD 者对非言语刺激的时序知觉表现出了任务差异, 这可

能是由于研究中 SOA 取值的不同造成的。如 Bebko 等(2006)的研究中 SOA 为 3 s, 而 de Boer-Schellekens 和 Eussen 等(2013)的研究中则为 40 ~ 320 ms。

4 视听时间整合缺陷的理论解释

4.1 时间绑定缺陷假说

时间绑定缺陷假说(the temporal binding deficit hypothesis)认为, ASD 者的时间加工缺陷是由于其神经激活的时间异常造成的, 主要表现为大脑的神经放电在时间上缺乏可靠性, 如过度超同步化放电、神经元突发性异常放电等(Brock, Brown, Boucher, & Rippon, 2002)。其假设的前提是, 大脑是通过特征绑定进行信息加工的。在视听整合过程中, 分散于不同皮层的信息组合的关键是不同特征神经激活之间的时间关系。当一簇视听信息流在相近的时间内输入时, 会激活大量神经元, 这些神经激活之间的时间关系发生紊乱会造成知觉绑定缺陷。例如, 有研究发现, ASD 者可以整合视听刺激, 但所需加工时间更长, 表现出 TBW 宽的特点(Foss-Feig et al., 2010; Kwakye et al., 2011)。这是由于其局部神经网络之间的时间绑定可以进行但功能较弱, 而局部神经网络内的时间绑定是完好甚至高于正常被试的。该假说是目前解释 ASD 者视听时间加工缺陷较核心的理论, 但由于其提出时间较早, 在有关特征捆绑的信息加工模式、时间绑定理论的有效性、对复杂刺激绑定模式的解释及 γ 波同步性激活等方面还存在不足(Brock et al., 2002)。

4.2 信噪比减少假说

信噪比减少假说(decreased signal-to-noise ratio hypothesis)认为, ASD 者神经编码的信噪比较低, 神经系统自发放电干扰和反应变异性较大, 导致其稳定性差而减弱了信息提取能力。视听信息在一定的时间范围内呈现时会相互影响, 当单通道神经激活的时间变异增大时, 会干扰双通道的整合加工(Rubenstein & Merzenich, 2003; Dinstein et al., 2012)。通常, 当呈现一个单通道刺激时, 刺激出现的瞬间便会激活与之相关的神经, 但 ASD 者对两者之间时间精确性的知觉较弱。例如, 有关研究探讨了单通道时间加工异常是否会造成 ASD 者的视听时间整合缺陷。结果发现, 与正常被试相比, 在视觉时间加工方面没有显著差异, 但其听觉时间加工敏感性显著低于正常组(Kwakye et al., 2011)。有关事件相关电位的研究表明, ASD 者在无噪音背景下对听觉刺激的反应与控制组在有噪音背景下的反应结果一样, 这表明其在基线水平下对听觉刺激的反应减弱(Russo, Zecker, Trommer, Chen, & Kraus, 2009)。该假说从神经机制角度对 ASD 者 TBW 宽的特点进行了解释, 但与时间绑定缺陷假说有所不同。前者认为 ASD 者主要是神经网络之间的时间加工存在异常, 而局部网络内的加工是完好的; 后者则强调是单通道的时间加工异常引起了双通道整合缺陷。

4.3 贝叶斯理论

该理论认为, 知觉是感觉刺激与内部表征最优匹配的结果, ASD 者在进行信息加工时过度依赖当前感觉信息, 而忽略先前经验(Pellicano & Burr, 2012)。在进行视听时间再校准时, 个体需要合理优化内部表征和实际感觉信号之间不同步的时间关系, 最小化同一事件感觉线索之间的差异, 促进视听整合。对正常人而言, 当两者之间存在差异时, 个体会灵活调整这种不符, 但由于 ASD 者感觉线索权重的增加, 其内在期望和实际感觉信号之间的不符会减少, 使得再校准能力减弱。内在表征和感觉噪音期望分布之间的强度差异是两者之间不符的主要因素, 这种差异会随着刺激复杂性的不同而变化。例如, 研究发现, ASD 者对言语刺激的快速视听时间再校准与正常组无差异, 而对非言语刺激的再校准能力显著减弱(Turi et al., 2016; Noel et al., 2017)。这是由于, 与言语刺激相比, 低水平刺激的感觉噪音更小。当呈现同样的感觉信号时, ASD 者对低水平刺激内部表征的噪音减少, 误差信号减弱, 而不能进行再校准。但言语刺激因其内部表征的噪音较大, 使得 ASD 组和控制组都可以进行。该理论解释了先前经验在视听时间整合中的作用, 但主要针对 ASD 者的快速视听时间再校准缺陷, 且目前该方面的实证研究主要集中于即时的快速再校准, 很少涉及建立在长时适应基础上的再校准。

4.4 预测编码假说

预测编码假说(predictive coding hypothesis)认为, 多感觉整合包括视听信息同时输入、与内部表征匹配、作出辨别或反应三个过程。在此过程中, 视听刺激会依据彼此之间的时间关系相互影响, 若视觉通道先完成加工, 便会为听觉信息

加工提供可靠的线索，进而提高加工效率。如当发出声音“ba”时，既可以利用听觉线索知觉，也可通过视觉线索知觉，如果先激活听觉通道，它就会促进视觉加工(Altieri, 2014)。由于 ASD 个体对环境信息的统计理解能力较弱，不能对视听刺激之间的时间关系进行精准把握。因此，当视听刺激的阈限都在一定的时间内激活时，其信息加工是彼此独立的，听觉不能对视觉产生影响(Pellicano & Burr, 2012)。例如，在“pop-out”效应中，ASD 者不能利用听觉线索促进视觉加工，致使视觉搜索效率没有提高(Collignon et al., 2013)。该理论从神经计算角度解释了 ASD 者的内隐时间加工缺陷，但目前的实证研究还很少，且大都是从间接角度进行验证，还需要直接证据的支撑。

综上所述，目前的理论从神经机制的异常激活模式和认知加工差异的层面对 ASD 者的视听时间整合缺陷进行了解释，得到了一些行为研究的支持，但相关的神经生理学研究还很少。此外，视听时间整合不仅是对刺激间物理时间量的加工，还涉及心理时间量加工，其整合过程中的神经激活存在时间进程，ASD 者和正常被试的激活规律是否相同，目前的理论还难以解释。未来需加强相关的神经机制研究，可以构建视听时间整合加工的理论模型，并对行为数据加以验证。例如，在有关视听言语整合的研究中，学者使用神经生理学技术构建了视听言语整合的神经加工模型，在正常被试群体中建立常模，比较其与 ASD 者的加工差异。结果发现，ASD 者视听言语整合缺陷不是由于其突触连接和可塑性弱，而是注意和多感觉加工缺陷导致的(Cuppini et al., 2017)。

5 小结与展望

视听时间整合探讨了视听整合中的时间加工机制。ASD 者的相关研究发现了其在视听整合的时间窗口、跨通道刺激的时序辨别、内隐时间线索加工等方面存在缺陷。研究中使用方法多样，但主要集中于行为层面，关注的焦点是其 TBW 异常，并通过使用不同刺激类型研究其与视听言语加工的关系。相关理论从神经加工机制、先前经验作用和多感觉加工过程等角度解释了其缺陷。不过，该领域仍存在一些问题亟待研究。

5.1 提高研究生态效度

ASD 者视听时间整合的研究方法目前主要包括用于外显时序知觉研究的 SJ、TOJ、优先注视任务以及基于内隐时间加工研究的声音诱发闪光错觉和“pip-pop”任务，但这些方法还存在不足。首先，进一步分离内隐研究中的注意因素和时间加工机制。在视听整合的研究中，声音的呈现会影响视觉加工，当视听刺激同时呈现时会产生整合效应，但当声音先于视觉刺激如提前 200 ms 呈现时，视觉加工效率的提高可能是由于声音引起的警觉作用而非整合(de Boer-Schellekens, Keetels, Eussen, & Vroomen, 2013)。因此 ASD 者视听整合缺陷是由于其不能精确知觉跨通道刺激间的时间线索，还是注意缺陷造成的，两者加工机制的相互作用值得深入探讨。其次，改进研究方法以提高对被试的适用性。ASD 是一组谱系障碍，包括很多亚型，症状差异和个体差异极大。目前的任务各具针对性，如 SJ、TOJ、“pip-pop”及声音诱发闪光错觉任务主要适用于高功能 ASD 被试，甚至有些高功能 ASD 者也无法完成，而优先注视任务主要用于幼儿的研究。未来需简化研究任务，进一步开发适用于低功能及不同亚型被试的任务。

5.2 整合理论解释

目前对 ASD 者视听时间整合缺陷的理论解释还缺乏普适性，如时间绑定缺陷假说和信噪比减少假说主要用于解释其 TBW 宽缺陷、贝叶斯理论主要对快速视听时间再校准不足进行了解释、预测编码理论解释了其内隐时间加工缺陷。但视听时间整合是一个动态的整体加工过程，可能存在多种加工方式。例如，有研究提出正常人对视听刺激的同步性知觉存在 4 种加工模式：大脑可以直接忽略视听刺激间的时间差；中枢神经可以对时间延迟的刺激进行补偿加工；个体会灵活调整视听刺激的不同时性使其差异最小化进而促进整合；对一个通道刺激的时间知觉会向另一个通道偏移(Vroomen & Keetels, 2010)。因此，有必要进一步整合相关研究，提出较全面、系统的理论解释。

5.3 精确量化诊断指标

时间窗、主观同时点和最小可觉差是目前衡量视听时间整合的重要指标。它们都可以进行精确量化，但要作为 ASD 者视听时间整合缺陷的诊断依据还需谨慎。首先，重视 TBW 宽度的量化。正常人的 TBW 一般为 300 ms，而研究发现 ASD 者为 600 ms (Foss-Feig et al., 2010)。但由于研究

中使用任务的差异、被试反应偏差的影响、刺激材料选取的不同以及统计标准差异等因素，不同研究中对 TBW 的计算还不够精确统一(Stevenson & Wallace, 2013)，如果未来可以证实其跨任务、跨年龄等的一致性，这将会为 ASD 的诊断提供重要依据；其次，综合使用三者。除了 TBW 宽度的量化，正常人适应视听刺激的时间差后 PSS 的偏移量表现为先增大后减小的趋势，峰值出现在刺激延迟 100~200 ms 左右，且不同个体时间知觉的 JND 也存在差异(Vroomen, Keetels, De Gelder, & Bertelson, 2004; Fujisaki, Shimojo, Kashino, & Nishida, 2004; Stevenson, Zemtsov, et al., 2012)。未来可以结合这三个指标，提高诊断依据的可靠性。

5.4 开发有效干预策略

对 ASD 者视听时间整合的干预不仅可以提高其时间知觉能力，促进视听整合，还可以进一步改善其言语知觉和社交缺陷。目前对 ASD 者视听时间整合的干预研究还比较少，而正常人已在行为和神经机制方面形成了两类较系统的干预模式：一种是自下而上再校准(bottom-up recalibration through exposure)，即通过从简单到复杂的任务训练逐步提高视听时间知觉敏感性；另一种是自上而下的反馈训练(top-down feedback training)，即通过反馈不断调整对视听刺激的时间知觉(Fujisaki et al., 2004; Powers, Hillock, & Wallace, 2009; Powers, Hevey, & Wallace, 2012)。干预中主要使用知觉学习任务，这种任务简单、易操作，可以适用于 ASD 被试。此外，电子游戏经验可以促进个体的时序知觉，研究表明，电子游戏经验丰富的个体在 SJ 和 TOJ 任务中视听时序辨别的敏感性更高(Donohue, Woldorff, & Mitroff, 2010)。未来可以借助多媒体技术对知觉学习任务进行改进，开发适用于 ASD 者的干预策略，这种干预方式不仅有趣、可操作性强，可以让被试在娱乐中提升视听时间知觉能力，还能够根据被试的年龄特征和功能症状提出有针对性的干预方案。

参考文献

Altieri, N. (2014). Multisensory integration, learning, and the predictive coding hypothesis. *Frontiers in Psychology, 5*(2), 257.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)* (5th ed.). Washington, DC: American Psychiatric Publishing.

Baum, S. H., Stevenson, R. A., & Wallace, M. T. (2015). Testing sensory and multisensory function in children with autism spectrum disorder. *Journal of Visualized Experiments,* (98), e52677.

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the Mcgurk effect. *Journal of Neuroscience, 30*(7), 2414–2417.

Bebko, J. M., Weiss, J. A., Demark, J. L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *Journal of Child Psychology and Psychiatry, 47*(1), 88–98.

Binder, M. (2015). Neural correlates of audiovisual temporal processing–Comparison of temporal order and simultaneity judgments. *Neuroscience, 300*, 432–447.

Brock, J., Brown, C. C., Boucher, J., & Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Development and Psychopathology, 14*(2), 209–224.

Collignon, O., Charbonneau, G., Peters, F., Nassim, M., Lassonde, M., Lepore, F.,... Bertone, A. (2013). Reduced multisensory facilitation in persons with autism. *Cortex, 49*(6), 1704–1710.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience, 16*(6), 1000–1009.

Cuppini, C., Ursino, M., Magosso, E., Ross, L. A., Foxe, J. F., & Molholm, S. (2017). A computational analysis of neural mechanisms underlying the maturation of multisensory speech integration in neurotypical children and those on the autism spectrum. *Frontiers in Human Neuroscience, 11*, 518.

de Boer-Schellekens, L., Eussen, M., & Vroomen, J. (2013). Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Frontiers in Integrative Neuroscience, 7*, 8.

de Boer-Schellekens, L., Keetels, M., Eussen, M., & Vroomen, J. (2013). No evidence for impaired multisensory integration of low-level audiovisual stimuli in adolescents and young adults with autism spectrum disorders. *Neuropsychologia, 51*(14), 3004–3013.

Dinstein, I., Heeger, D. J., Lorenzi, L., Minshew, N. J., Malach, R., & Behrmann, M. (2012). Unreliable evoked responses in autism. *Neuron, 75*(6), 981–991.

Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics, 72*(4), 1120–1129.

Doyle-Thomas, K. A. R., Goldberg, J., Szatmari, P., & Hall, G. B. C. (2013). Neurofunctional underpinnings of audiovisual emotion processing in teens with autism spectrum disorders. *Frontiers in Psychiatry, 4*, 48.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An

extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research, 203*(2), 381–389.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. Y. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*(7), 773–778.

Glessner, J. T., Wang, K., Cai, G. Q., Korvatska, O., Kim, C. E., Wood, S.,... Hakonarson, H. (2009). Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature, 459*(7246), 569–573.

Greenfield, K., Ropar, D., Smith, A. D., Carey, M., & Newport, R. (2015). Visuo-tactile integration in autism: Atypical temporal binding may underlie greater reliance on proprioceptive information. *Molecular Autism, 6*(1), 51.

Grossman, R. B., Steinhart, E., Mitchell, T., & McIlvane, W. (2015). "Look who's talking!" Gaze patterns for implicit and explicit audio-visual speech synchrony detection in children with high-functioning autism. *Autism Research, 8*(3), 307–316.

Hairston, W. D., Burdette, J. H., Flowers, D. L., Wood, F. B., & Wallace, M. T. (2005). Altered temporal profile of visual–auditory multisensory interactions in dyslexia. *Experimental Brain Research, 166*(3-4), 474–480.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(1), 5–25.

Hillock, A. R., Powers, A. R., & Wallace, M. T. (2011). Binding of sights and sounds: Age-related changes in multisensory temporal processing. *Neuropsychologia, 49*(3), 461–467.

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science, 15*(5), 688–696.

Hocking, J., & Price, C. J. (2008). The role of the posterior superior temporal sulcus in audiovisual processing. *Cerebral Cortex, 18*(10), 2439–2449.

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., & Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in Integrative Neuroscience, 4*, 129.

Laasonen, M., Service, E., & Virsu, V. J. (2001). Temporal order and processing acuity of visual, auditory, and tactile perception in developmentally dyslexic young adults. *Cognitive, Affective, & Behavioral Neuroscience, 1*(4), 394–410.

Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child Development, 85*(2), 685–694.

Megnin, O., Flitton, A., Jones, C. R. G., de Haan, M., Baldeweg, T., & Charman, T. (2012). Audiovisual speech integration in autism spectrum disorders: ERP evidence for atypicalities in lexical-semantic processing. *Autism Research, 5*(1), 39–48.

Noel, J. P., De Niear, M. A., Stevenson, R., Alais, D., & Wallace, M. T. (2017). Atypical rapid audio-visual temporal recalibration in autism spectrum disorders. *Autism Research, 10*(1), 121–129.

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience & Biobehavioral Reviews, 36*(2), 836–854.

Patten, E., Watson, L. R., & Baranek, G. T. (2014). Temporal synchrony detection and associations with language in young children with ASD. *Autism Research and Treatment, 2014*, Article ID 678346.

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16*(10), 504–510.

Powers, A. R., Hevey, M. A., & Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263–6274.

Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *Journal of Neuroscience, 29*(39), 12265–12274.

Rippon, G., Brock, J., Brown, C., & Boucher, J. (2007). Disordered connectivity in the autistic brain: Challenges for the 'new psychophysiology'. *International Journal of Psychophysiology, 63*(2), 164–172.

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior, 2*(5), 255–267.

Russo, N., Foxe, J. J., Brandwein, A. B., Altschuler, T., Gomes, H., & Molholm, S. (2010). Multisensory processing in children with autism: High-density electrical mapping of auditory–somatosensory integration. *Autism Research, 3*(5), 253–267.

Russo, N., Zecker, S., Trommer, B., Chen, J. L., & Kraus, N. (2009). Effects of background noise on cortical encoding of speech in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(8), 1185–1196.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*(1), 147–152.

Shi, Z. H., Chen, L. H., & Müller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: The influence of time interval. *Experimental Brain Research, 203*(4), 723–735.

Stevenson, R. A., Fister, J. K., Barnett, Z. P., Nidiffer, A. R., & Wallace, M. T. (2012). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Experimental Brain Research, 219*(1), 121–137.

Stevenson, R. A., Segers, M., Ferber, S., Barense, M. D., Camarata, S., & Wallace, M. T. (2016). Keeping time in the brain: Autism spectrum disorder and audiovisual temporal processing. *Autism Research, 9*(7), 720–738.

Stevenson, R. A., Siemann, J. K., Brown, S. T., Woynaroski, T. G., Segers, M., Bebko, J., & Wallace, M. (2013). Atypical multisensory integration in Autism Spectrum

Disorders: Cascading impacts of altered temporal processing. *Multisensory Research, 26*, 25.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience, 34*(3), 691–697.

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., & Wallace, M. T. (2014). Evidence for diminished multisensory integration in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(12), 3161–3167.

Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., & James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *NeuroImage, 55*(3), 1339–1345.

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research, 227*(2), 249–261.

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1517–1529.

Turi, M., Karaminis, T., Pellicano, E., & Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum. *Scientific Reports, 6*, 21756.

van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *Journal of Neuroscience, 33*(37), 14633–14637.

van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1053–1065.

Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain Research, 1111*(1), 134–142.

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research, 22*(1), 32–35.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics, 72*(4), 871–884.

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105–123.

Williams, E. L., & Casanova, M. F. (2010). Autism and dyslexia: A spectrum of cognitive styles as defined by minicolumnar morphometry. *Medical Hypotheses, 74*(1), 59–62.

Audiovisual temporal integration in autism spectrum disorders

LI Taotao[1]; HU Jinsheng[1]; WANG Qi[1]; LI Chengshi[1]; LI Songze[1];
HE Jianqing[1]; LI Chenyang[1]; LIU Shuqing[2]

([1] *Department of Psychology, Liaoning Normal University, Dalian* 116029, *China*)

([2] *College of Basic Medical Sciences, Dalian Medical University, Dalian 116044, China*)

Abstract: Audiovisual temporal integration is the fact that the relative timing of audio-visual stimulation should be bound into a unified percept. The audiovisual temporal integration of autism spectrum disorders (ASD) was confirmed to be impaired mainly in four aspects. *First*, the temporal binding window is wider and more symmetrical. *Second*, ASD got weaker ability of rapid audiovisual temporal recalibration. *Third*, the effect that visual target search can be improved by a synchronous sound decreased has been found to be missing in ASD in some studies. *Last*, ASD showed impaired audiovisual temporal acuity for speech stimuli. Various methods are used to probe for temporal aspects of audiovisual integration, such as sound-induced flash illusion task and "pip-pop" task. The simultaneity judgment task, temporal order judgment task and preferential-looking task are mainly used to measure intersensory synchrony. Relevant theories from abnormal neural processing, the internal priors and the interaction between multiple sensory modals explain those defects. In future, researchers should improve the ecological validity of study, further integrate different theoretical explanations and accurate quantitative diagnostic indicators, as well as exploit the effective intervention strategies.

Key words: autism spectrum disorders; audiovisual temporal integration; temporal binding window; temporal synchrony